

Analysis of a Marine Picoplankton Community by 16S rRNA Gene Cloning and Sequencing

THOMAS M. SCHMIDT,[†] EDWARD F. DELONG,[‡] AND NORMAN R. PACE*

Department of Biology and Institute for Molecular and Cellular Biology, Indiana University, Bloomington, Indiana 47405

Received 7 January 1991/Accepted 13 May 1991

The phylogenetic diversity of an oligotrophic marine picoplankton community was examined by analyzing the sequences of cloned ribosomal genes. This strategy does not rely on cultivation of the resident microorganisms. Bulk genomic DNA was isolated from picoplankton collected in the north central Pacific Ocean by tangential flow filtration. The mixed-population DNA was fragmented, size fractionated, and cloned into bacteriophage lambda. Thirty-eight clones containing 16S rRNA genes were identified in a screen of 3.2×10^4 recombinant phage, and portions of the rRNA gene were amplified by polymerase chain reaction and sequenced. The resulting sequences were used to establish the identities of the picoplankton by comparison with an established data base of rRNA sequences. Fifteen unique eubacterial sequences were obtained, including four from cyanobacteria and eleven from proteobacteria. A single eucaryote related to dinoflagellates was identified; no archaeobacterial sequences were detected. The cyanobacterial sequences are all closely related to sequences from cultivated marine *Synechococcus* strains and with cyanobacterial sequences obtained from the Atlantic Ocean (Sargasso Sea). Several sequences were related to common marine isolates of the γ subdivision of proteobacteria. In addition to sequences closely related to those of described bacteria, sequences were obtained from two phylogenetic groups of organisms that are not closely related to any known rRNA sequences from cultivated organisms. Both of these novel phylogenetic clusters are proteobacteria, one group within the α subdivision and the other distinct from known proteobacterial subdivisions. The rRNA sequences of the α -related group are nearly identical to those of some Sargasso Sea picoplankton, suggesting a global distribution of these organisms.

Marine picoplankton, organisms between 0.2 and 2 μ m in diameter (18), are thought to play a significant role in global mineral cycles (2), yet little is known about the organismal composition of picoplankton communities. Standard microbiological techniques that involve the study of pure cultures of microorganisms provide a glimpse at the diversity of picoplankton, but most viable picoplankton resist cultivation (8, 10). The inability to cultivate organisms seen in the environment is a common bane of microbial ecology (1).

As an alternative to reliance on cultivation, molecular approaches based on phylogenetic analyses of rRNA sequences have been used to determine the species composition of microbial communities (14). The sequences of rRNAs (or their genes) from naturally occurring organisms are compared with known rRNA sequences by using techniques of molecular phylogeny. Some properties of an otherwise unknown organism can be inferred on the basis of the properties of its known relatives because representatives of particular phylogenetic groups are expected to share properties common to that group. The same sequence variations that are the basis of the phylogenetic analysis can be used to identify and quantify organisms in the environment by hybridization with organism-specific probes.

Approaches that have been used to obtain rRNA sequences, and thereby identify microorganisms in natural samples without the requirement of laboratory cultivation, include direct sequencing of extracted 5S rRNAs (19, 20),

analysis of cDNA libraries of 16S rRNAs (21, 24), and analysis of cloned 16S rRNA genes obtained by amplification using the polymerase chain reaction (PCR) (3). Each of these approaches potentially imposes a selection on the sequences that are analyzed. Minor constituents of communities may not be detected using the 5S rRNA because only abundant rRNAs can be analyzed. Methods that copy naturally occurring sequences in vitro before cloning potentially select for sequences that interact particularly favorably with primers or polymerases.

In this study we have characterized 16S rRNA sequences from a Pacific Ocean picoplankton population by using methods chosen to minimize selection of particular sequences. DNA from the mixed population was cloned directly into phage λ , and rRNA gene-containing clones were identified subsequently. The methods used in the analysis are applicable to other natural microbial communities. The results are correlated with sequences from cultivated organisms and with 16S rRNA gene sequences selected by PCR from Atlantic picoplankton (3).

MATERIALS AND METHODS

Collection and microscopy of picoplankton. Picoplankton were collected in the north central Pacific Ocean at the ALOHA Global Ocean Flux Study site (22°45'N, 158°00'W) from aboard the R/V *Moana Wave*. As previously detailed (5), seawater was pumped onboard through a 10- μ m-pore-size Nytex filter and concentrated by tangential flow filtration using 10 ft² (ca. 6,450 cm²) of 0.1 μ m-pore-size fluorocarbon membrane, and cells were pelleted from the resulting concentrate by centrifugation and frozen. Sampling began on

* Corresponding author.

[†] Present address: Department of Microbiology, Miami University, Oxford, OH 45056.

[‡] Present address: Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

1 December 1988 and continued through 3 December 1988 within a 6-mile (ca. 9.7-km) radius of the ALOHA station.

For electron microscopy, concentrated picoplankton were fixed in 1% glutaraldehyde, pelleted by centrifugation, and suspended in TE buffer (10 mM Tris-HCl [pH 7.4], 1 mM EDTA) containing 0.05% Nonidet P-40 (Sigma Chemical Co., St. Louis, Mo.) and 10 mM NaCl. The suspended cells were allowed to settle overnight in a humidified chamber onto glass coverslips coated with Cel-Tak tissue adhesive (Biopolymers, Inc., Farmington, Conn.). The coverslips were washed in distilled water, serially transferred through 25, 50, and 70% ethanol, critical point dried, mounted on scanning electron microscope stubs and sputter-coated with gold-palladium. Samples were viewed in a Cambridge 5250 MK2 scanning electron microscope.

Purification and cloning of mixed-population DNA. Cell pellets were thawed on ice in 4.5 ml of 40 mM EDTA–0.75 M sucrose–50 mM Tris-HCl, pH 8.3. Lysozyme was added to a final concentration of 1 mg/ml, and the suspension was incubated at 37°C for 30 min. After the addition of 800 µg of proteinase K and sodium dodecyl sulfate (SDS) to a final concentration of 0.5% (wt/vol), the mixture was incubated for an additional 2 h at 37°C. Phase microscopy indicated that cell lysis was complete. Polysaccharides and residual proteins were aggregated by addition of hexadecyltrimethyl ammonium bromide to a final concentration of 1.0% (wt/vol) in the presence of 0.7 M sodium chloride and a 20-min incubation at 65°C. The protein and polysaccharide complexes were removed by extraction with an equal volume of chloroform-isoamyl alcohol (24:1), followed by extraction with phenol-chloroform-isoamyl alcohol (50:49:1). Nucleic acids were recovered by the addition of 0.6 volume of isopropanol and centrifugation for 10 min at 10,000 rpm in an HB-4 rotor.

The pellet was suspended to a concentration of between 50 and 100 µg/ml in TE buffer. Cesium chloride was added to a final concentration of 1.075 g/ml in 5-ml polyallomer tubes, each containing 1 mg of ethidium bromide. A density gradient was established by centrifugation at 55,000 rpm for 16 h in a VTi 65 rotor. Genomic DNA was visualized by ethidium bromide fluorescence, and a broad zone centered about the DNA band was withdrawn with an 18-gauge hypodermic needle. Ethidium bromide was extracted from the DNA with water-saturated butanol, and cesium chloride was removed by dialysis against TE buffer, pH 7.8.

The purified, mixed-population DNA was partially digested with the restriction endonuclease *Sau3A*, and DNA fragments ranging in size from 10 to 20 kb were recovered from a 5 to 25% NaCl velocity gradient spun at 37,000 rpm in a Beckman SW41 rotor for 4.5 h at 25°C as detailed by Kaiser and Murray (9). Bacteriophage λ cloning vector EMBL3 was prepared by digestion with *Bam*HI and *Eco*RI and removal of the stuffer fragment by selective precipitation (Stratagene, La Jolla, Calif.). The size-fractionated fragments of picoplankton DNA were ligated to the insertion vector with T4 DNA ligase and then packaged in vitro with Gigapack Gold packaging extracts (Stratagene) as recommended by the manufacturer. The titer of the library was determined by mixing aliquots of the packaged DNA with *E. coli* host strain P2 392.

Identification of rRNA gene-containing recombinants. Aliquots of the recombinant library were mixed with *E. coli* host strain LE 392 (Stratagene) such that 1×10^4 to 5×10^4 plaques developed per 150-mm plate. Plaques were transferred to nitrocellulose (Schleicher & Schuell, Keene, N.H.), denatured with 0.5 M NaOH in 1.5 M NaCl, neutral-

ized in 0.5 M Tris-HCl, pH 7.5, in 1.5 M NaCl, and baked at 80°C for 2 h under vacuum.

Nitrocellulose filters were incubated in hybridization buffer (4) for 15 to 30 min at 42°C before the addition of approximately 10^7 cpm of either the mixed-kingdom or oligodeoxynucleotide probe (see below). When the mixed-kingdom probe was used, hybridization reactions were incubated at 65°C for 5 h and then slowly cooled to 42°C in an overnight incubation. Filters were washed once in SET (150 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl, pH 7.8) for 10 min at room temperature, followed by two successive washes in SET for 10 min at 45°C. Hybridization with the oligonucleotide probe was performed at 37°C overnight, followed by one 10-min wash in SET at room temperature and two successive 10-min washes in SET at 37°C. After drying, filters were exposed to X-ray film for 4 to 48 h.

A mixed-kingdom probe (14) was prepared from 16S-like rRNAs purified from *Oceanospirillum linum*, *Sulfolobus solfataricus*, and *Saccharomyces cerevisiae*. The respective small-subunit RNAs were purified by polyacrylamide gel electrophoresis, electroeluted from the gel, and partially hydrolyzed by incubation for 15 min in 100 mM NaHCO₃-Na₂CO₃, pH 9.0, at 90°C. The RNA was precipitated in 0.3 M sodium acetate and 2 volumes of ethanol. The pellet was washed with 70% ethanol to remove precipitated salts and suspended at 1.0 µg/ml in 10 mM Tris-HCl, pH 7.6. Phage T4 polynucleotide kinase (Pharmacia, Piscataway, N.J.) was used to 5' end label the RNA fragments (17) with 0.5 mCi of [γ -³²P]ATP. Labeled probe was purified on Sephadex G-50 columns, using a running buffer of 100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, and 0.1% SDS. The purified, end-labeled probe had a specific activity of 10^7 to 10^8 cpm/µg of RNA.

The oligonucleotide hybridization probe 515F (GTGCCA GCMGCCGCGG), identical to a universally conserved region of the small-subunit rRNA (11), was synthesized on an Applied Biosystems automated DNA synthesizer and purified by polyacrylamide gel electrophoresis. The probe was 5' end labeled as described above and then purified on a C8 reverse-phase Bond Elut column (Analytichem International, Harbour City, Calif.) as described previously (4).

Plaques hybridizing with rRNA-specific probes were purified by two subsequent rounds of isolation. Phage DNA was purified from plate lysates on Prep-Eze columns (5'-3', West Chester, Pa.) as described previously (6).

Plaque dot assay. A plaque dot assay (15) was used to sort clones at the kingdom level so that appropriate primers could be selected for PCR amplification of the ribosomal DNA (rDNA) portion of the cloned fragments. To optimize the plaque dot assay for maximal phage production and to test the relative efficacy of filter papers used in the plaque lifts, 1-µl aliquots of dilutions of a purified, rRNA gene-containing phage preparation were spotted onto a fresh lawn of *Escherichia coli*. After incubation at 37°C for 24 h, the plaques were lifted by using either nitrocellulose or Hy-Bond filter paper (Amersham, Arlington Heights, Ill.) and hybridized with the eubacterial (*O. linum*) rRNA probe.

Amplification of rDNA genes. Eubacterial rDNA was amplified from λ clones by PCR (16), using amplification primers specific to nucleotide positions 50 to 68 (forward primer; AACACATGCAAGTCGAACG) and 536 to 519 (reverse primer; GWATTACCGCGGCKGCTG) of the *E. coli* 16S rRNA (11). Purified λ DNA (100 to 400 ng) was added to the amplification reaction mixture, which contained (in a 100-µl total volume) 10 µl of $10\times$ reaction buffer (100 mM Tris-HCl [pH 8.3], 500 mM KCl, 15 mM MgCl₂, 0.01%

[wt/vol] gelatin), 5 μ l of 1% Nonidet P-40 (Sigma), 1.5 mM each dATP, dGTP, dCTP, and dTTP, 0.5 μ g of each primer, and 2 U of *Taq* DNA polymerase. The reaction mixture was overlaid with mineral oil and incubated in a thermal cycler (Perkin Elmer-Cetus) as follows: 1.5 min of denaturation at 92°C, primer annealing at 37°C for 1.5 min, heating to 72°C at 2°C/s, and elongation for 2 min, which was extended for 5 s after each cycle. Following 20 rounds of amplification, the reaction mixture was extracted once with phenol-chloroform-isoamyl alcohol (50:49:1), and the amplified product was precipitated from 0.3 M ammonium acetate and 1 volume of isopropanol. The nucleic acid was pelleted by centrifugation for 15 min and washed once in 70% ethanol-TE. The pellet was redissolved in the amplification cocktail, and the amplification was repeated as described above except that only the forward primer was added to the reaction mixture. Following an additional 20 rounds of amplification, the reaction mix was extracted once with phenol/chloroform and precipitated twice with ammonium acetate and isopropanol as described above. The final pellet was suspended in 10 μ l of TE. Typically, 2 to 4 μ l of the resuspended pellet was used per sequencing reaction, using a 5'-³²P-labeled primer (above) and the Klenow fragment of *E. coli* DNA polymerase (14).

Sequence analysis. Sequences were aligned manually on a collection of rRNA sequences on the basis of conserved regions of sequence and secondary structure of the 16S rRNA (25). Regions of ambiguous alignment were omitted from subsequent analyses. The evolutionary distance between each pair of sequences was calculated, and a least squares method was used to infer the phylogenetic tree most consistent with the pairwise distance estimates (12).

Nucleotide sequences accession numbers. Sequences used in phylogenetic analyses are all available from the GenBank or EMBL data base. The new sequences are available under the following GenBank accession numbers: ALO 7, M64536; ALO 23, M64526; ALO 37, M64531; ALO 11, M64522; ALO 30, M64529; ALO 33, M64530; ALO 29, M64528; ALO 18, M64524; ALO 4, M64534; ALO 40, M64535; ALO 24, M64527; ALO 17, M64523; ALO 39, M64533; ALO 38, M64532; and ALO 21, M64525.

RESULTS

The approach used here to characterize marine picoplankton without cultivation and with minimum selectivity is summarized in Fig. 1, with some results of the study. Picoplankton sufficient for the analysis were concentrated from oligotrophic oceanic water by tangential flow filtration and centrifugation (5). As illustrated in Fig. 2, the cells collected were between 0.2 and 2 μ m in diameter and so are defined as picoplankton (18). The morphologies of the organisms in the population analyzed here are typical of marine picoplankton (2, 23). A total of 471 μ g of purified DNA was recovered as detailed in Materials and Methods from a 140-mg portion of the 560-mg picoplankton cell pellet. The extracted DNA was of high molecular size (>40 kb) and uniformly susceptible to digestion with restriction endonucleases. A library consisting of 10⁷ recombinants was constructed in phage λ EMBL3 from partially digested (*Sau*3A) and size-fractionated (10- to 20-kb) picoplankton DNA.

Isolation of rRNA gene-containing clones. Plaque hybridization was used to detect rRNA gene-containing recombinants. Since rRNA genes in the recombinant library were from unknown organisms, it was necessary for detection of those genes to rely upon hybridization probes that bind to

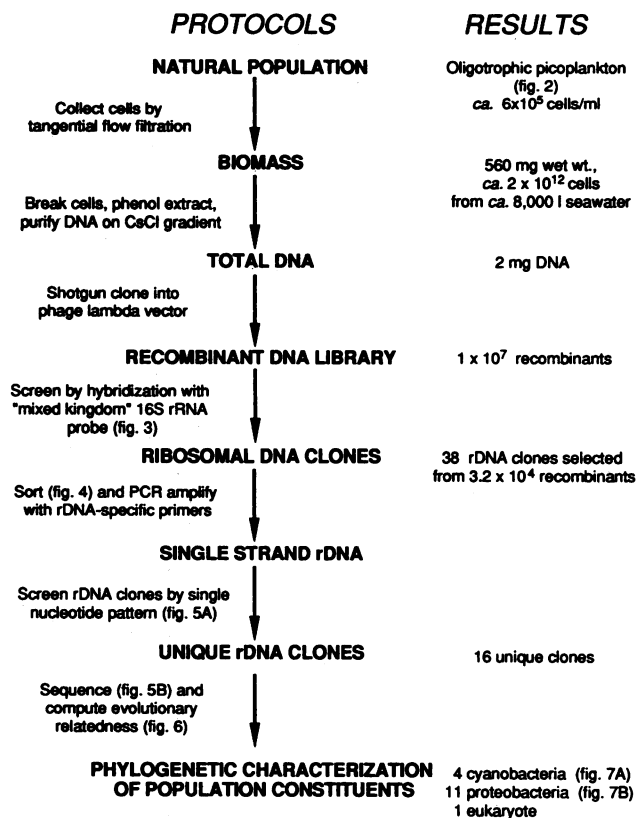


FIG. 1. Flow chart of the protocols used to characterize marine picoplankton without cultivation and a summary of some results.

generally conserved features of 16S rRNAs. We evaluated two types of hybridization probes for identifying the picoplankton genes. One, a mixed-kingdom probe, consisted of a mixture of partially hydrolyzed 16S rRNAs derived from one representative of each of the primary kingdoms (25): eubacteria (represented by *O. linum*), eucaryotes (*S. cerevisiae*), and archaebacteria (*S. solfataricus*). The extent of sequence conservation in 16S-like rRNAs is such that cross-species hybridization is readily detected among representatives of a particular kingdom (14). The probe mixture of rRNAs from each of the primary kingdoms should detect the rRNA genes of all or most life forms. The second hybridization probe that we tested was a 16-nucleotide sequencing primer identical to a universally conserved sequence in 16S rRNAs (Materials and Methods). This probe, too, in principle should detect all rRNA genes in the recombinant library. The hybridization of these two types of probes to duplicate, representative plaque lifts of the picoplankton library is shown in Fig. 3. All plaques identified with the mixed-kingdom probe (Fig. 3A) also were identified with the oligonucleotide probe (Fig. 3B). However, the oligonucleotide probe also bound to many additional plaques. Subsequent sequence analysis revealed that all recombinants identified by the mixed-kingdom probe contained rRNA genes, whereas none of several recombinants analyzed that bound the oligonucleotide probe, but not the mixed-kingdom probe, contained recognizable rRNA gene sequences. We therefore used and recommend the mixed-kingdom probe for hybridization screening for uncharacterized rRNA genes.

Thirty-eight clones containing rDNA inserts were identi-

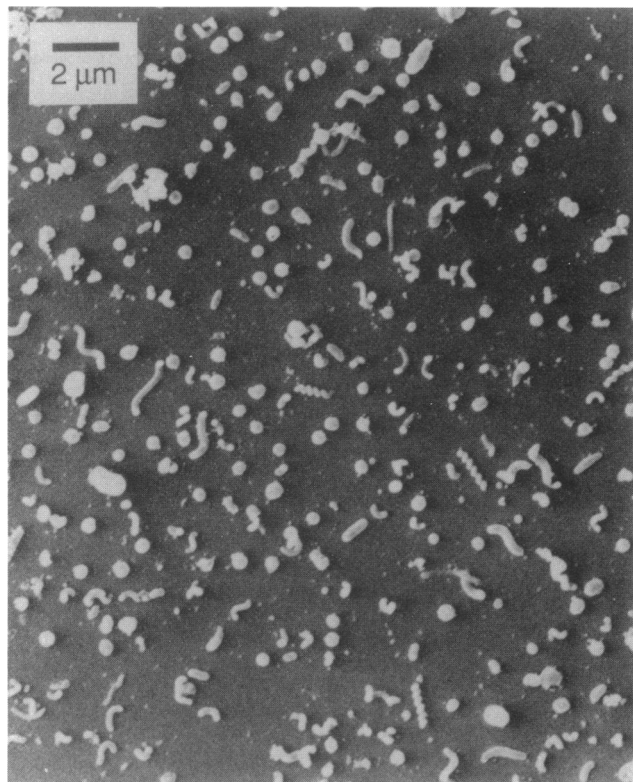


FIG. 2. Scanning electron micrograph of picoplankton represented in the clone library. Picoplankton concentrated from the ALOHA collection site were fixed and prepared for scanning electron microscopy as detailed in Materials and Methods.

fied among 3.2×10^4 recombinants with use of the mixed-kingdom probe and were purified by two successive rounds of single-plaque isolation. rDNA-containing phage then were characterized with rRNA probes in a plaque dot assay using kingdom-specific hybridization conditions in order to select PCR primers for preparing templates for sequence analysis. In this assay, phage are dotted onto a lawn of indicator host, and following plaque development, plaques are lifted onto a

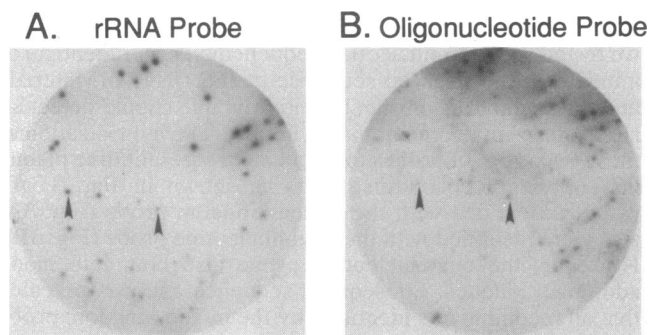


FIG. 3. Identification of rDNA-containing clones. Filter lifts of the recombinant bacteriophage library were probed as detailed in Materials and Methods with a mixed-kingdom rRNA probe consisting of 16S rRNAs from *O. linum* (eubacterium), *S. solfataricus* (archaeobacterium), and *S. cerevisiae* (eucaryote) or an oligonucleotide probe complementary to a universally conserved 16S rRNA sequence.

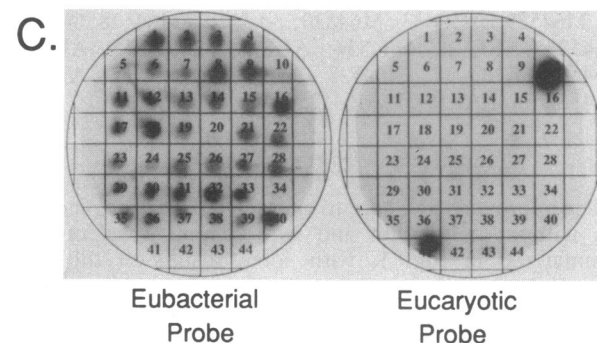
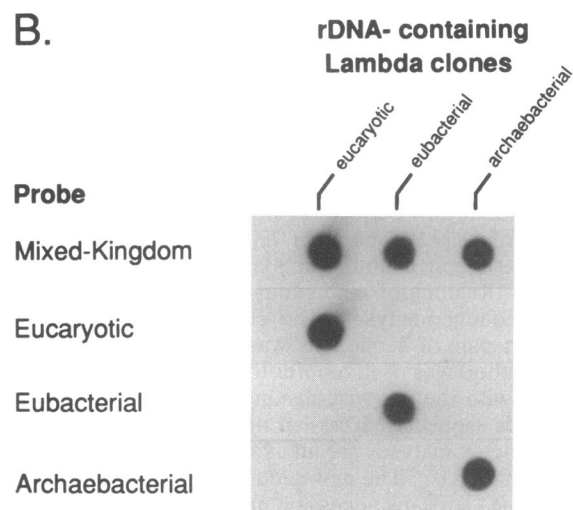
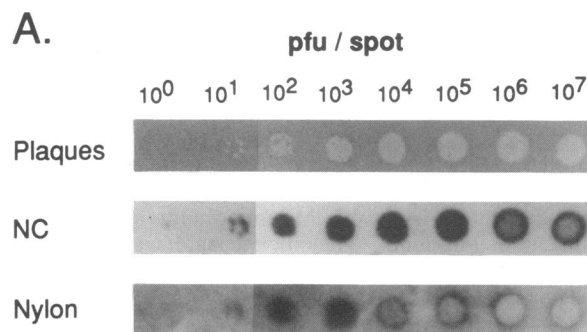


FIG. 4. Plaque dot assay. (A) Plaques of eubacterial (*Xenorhabdus luminescens*) 16S rDNA-containing bacteriophage, resulting from application of 10^7 phage per dot, were lifted onto either nitrocellulose (NC) or nylon (Hy-Bond) filters and hybridized with the eubacterium-specific (*O. linum*) rRNA probe. (B) Autoradiograms of nitrocellulose filters from four replicate plaque lifts, demonstrating kingdom specificity of the rRNA probes. The 16S rDNAs contained in the recombinant phage were as follows: eucaryotic, *S. cerevisiae*; eubacterial, *X. luminescens*; and archaeobacterial, *S. solfataricus*. The ^{32}P -labeled rRNA probes were as detailed in Materials and Methods. (C) Autoradiograms of plaque dot assays of purified, rDNA-containing bacteriophage from the picoplankton library probed with either the eubacterial or eucaryotic probe (B) as indicated. Bacteriophage containing the rDNA operon of *S. cerevisiae* were spotted at position 41.

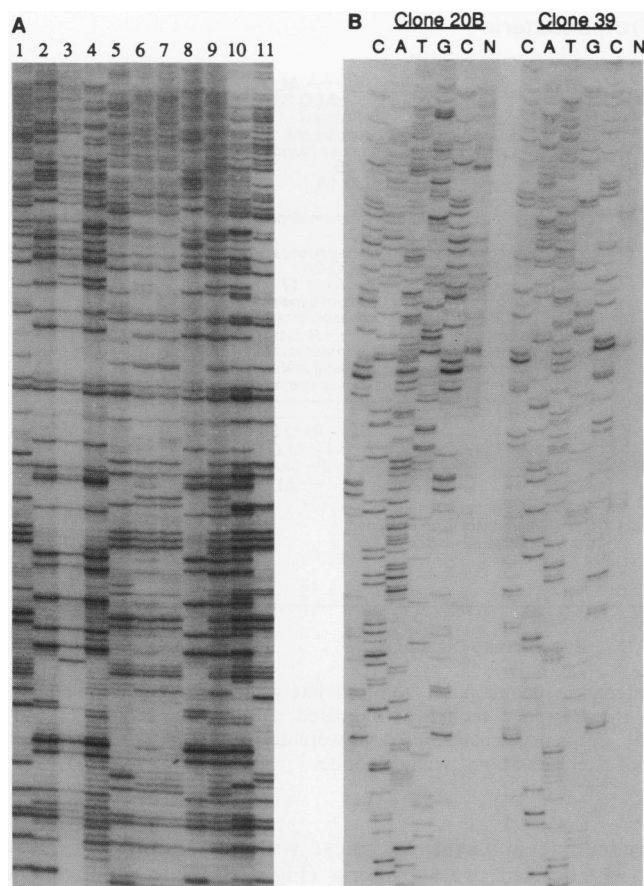


FIG. 5. Sorting and sequence analysis of rDNA-containing clones. (A) Single-nucleotide sequence pattern for 11 picoplankton clones, produced as detailed in Materials and Methods by using a PCR-amplified template, primer 519R, and a single dideoxynucleotide (ddA). (B) Sequence determination by dideoxynucleotide chain termination, using a PCR-generated, single-stranded template from two of the picoplankton clones and primer 519R.

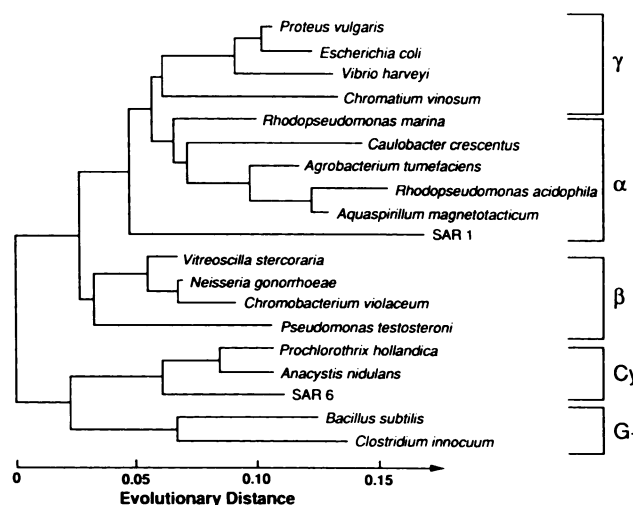
membrane filter for hybridization (15). As documented in Fig. 4, this assay is very sensitive to the phage concentration used as the dot inoculum and to the membrane filter to which plaques are adsorbed for the hybridization assay. The optimal titer for phage production in this assay is between 10^2 and 10^4 PFU/ μ l. At titers above 10^7 phage per ml, the signal intensity decreased, presumably due to poor phage production in hyperinfected cells (Fig. 4A). Additionally, the extent of hybridization was greater and background binding less when nitrocellulose filters were used for the plaque lifts in place of Hy-Bond filters (Fig. 4A).

The 38 purified phage preparations were diluted to titers of between 10^3 and 10^4 PFU/ μ l, and then 1- μ l aliquots were used in the plaque dot assay. The individual rRNAs that comprise the mixed-kingdom probe can be used independently under conditions of increased stringency to determine the kingdom affiliation of particular clones (Fig. 4B). Three successive plaque dot lifts of the picoplankton clones were hybridized individually with the three radiolabeled rRNAs that constitute the mixed-kingdom probe. A single clone containing eucaryotic rDNA was identified (Fig. 4C). The remaining 37 clones hybridized with the eubacterial-kingdom

probe (Fig. 4C). No archaeobacterial rDNA-containing recombinants were identified.

Sequence and phylogenetic analysis of picoplankton rDNA clones. A two-step, asymmetric amplification using PCR was used to generate ca. 500-nucleotide, single-stranded DNA for use as a template in sequencing. Amplification primers (Materials and Methods) were selected based on the kingdom affiliation of the clones. To distinguish unique and duplicate clones, sequencing reactions using the different templates were carried out with a single dideoxynucleotide (ddA) and sequencing primer 519R, and reaction products were analyzed in a denaturing 8% polyacrylamide gel (Fig. 5A). When clones shared identical locations of nucleotide A in the region sequenced, they were considered to be identical clones, and a representative was selected for detailed se-

A. Partial Mask



B. Full Mask

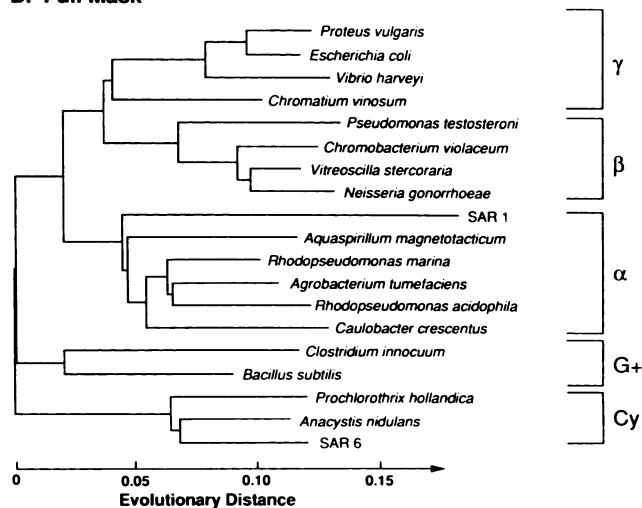


FIG. 6. Comparison of partial and full sequences for the computation of evolutionary relatedness of selected eubacteria. Phylogenetic trees were calculated (12) by using approximately 200 nucleotides (nucleotides ca. 300 to 500, *E. coli* numbering) (A) or the full sequences of the 16S rRNAs of the indicated organisms (B). The scale represents the number of fixed mutations per nucleotide.

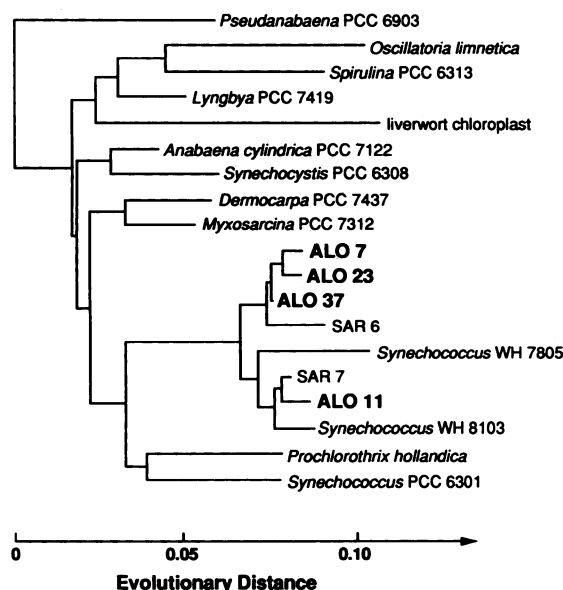
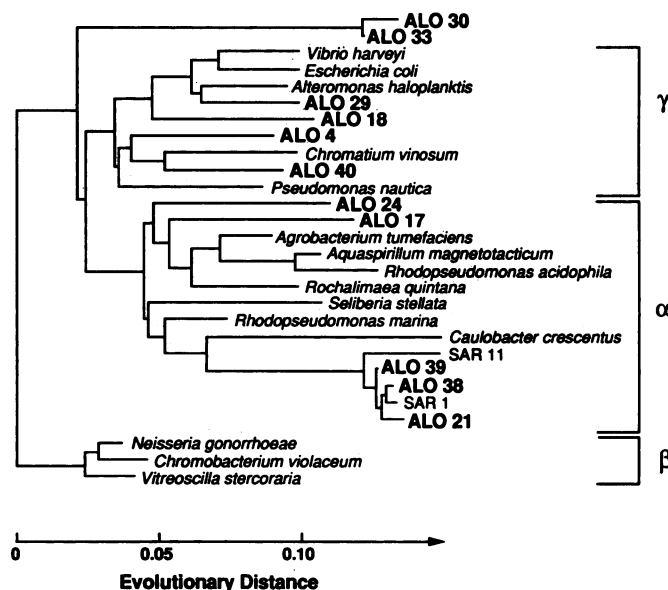
A. Cyanobacteria**B. Proteobacteria**

FIG. 7. Phylogenetic associations of cyanobacterial and proteobacterial marine picoplankton. Approximately 250-nucleotide sequences of the cloned Pacific Ocean picoplankton rRNA genes were aligned with the reference sequences indicated, and phylogenetic trees were constructed (12). Sequences designated ALO were determined in this study. Sequences designated SAR were derived by PCR from Sargasso Sea (Atlantic Ocean) picoplankton (3). The scale represents the number of fixed mutations per nucleotide.

quencing. Sixteen different rDNAs were identified among the 38 clones inspected.

The goal of this sequence analysis was to establish the phylogenetic affinities of the picoplankton contributing the rDNAs. The full sequence of the 16S rRNA genes is not necessary for phylogenetic assignment, as long as deeply diverging lineages are not at question. It has been shown previously that phylogenetic trees constructed by using different and limited portions of the 16S rRNA sequence are the same as those generated by using full sequences (11). We chose for focus a portion of the 16S rRNA gene that corresponds to nucleotides ca. 250 to 500 in the *E. coli* rRNA. This region of the molecule contains substantial sequence variability in conserved secondary structure and so is useful for rapid sequence comparisons. A phylogenetic tree constructed (Materials and Methods) with only 200 nucleotides of this region (Fig. 6A) is largely congruent with one calculated by using most of the 16S rRNA of the major phyla of bacteria considered (Fig. 6B). Some of the deeper-branching orders of the groups differ in the trees produced by using the different sequence extents, but the groups established are the same.

Approximately 250 nucleotides of sequence were determined from each unique PCR-amplified template, using ³²P-labeled 519R as the sequencing primer (Fig. 5B). The results of the phylogenetic analyses are summarized in Fig. 7.

All but one of the rDNA clones obtained were eubacterial, and they fell into only two of the approximately 12 major phylogenetic groups of eubacteria (25). About half of the inspected clones were representative of cyanobacteria, closely related or identical to one another, in a tight phylogenetic cluster (Fig. 7A) with two marine *Synechococcus* isolates (WH8103 and WH7805) and with sequences retrieved by PCR from Sargasso Sea picoplankton (3). The

remaining eubacterial clones, with few duplicates, were derived from proteobacteria (Fig. 7B). The phylogenetic analysis solidly establishes most of them with either the α or γ subdivision of proteobacteria. However, two of the sequences (ALO 30 and ALO 33), although clearly proteobacterial, fall outside the known subdivisions. This may be due to analytical inaccuracies that result from use of only 250 nucleotides in inferring the tree topology. It also is possible that ALO 30 and ALO 33 represent a new and deep divergence in the proteobacteria. More extensive sequence analysis will be required to resolve these relationships. All of the characterized sequences also exhibit signature sequences (26) exclusively characteristic of the phylogenetic groups to which the tree analysis assigns the sequences. The single eucaryotic rDNA clone recovered is most similar (90% identity) to the 16S rRNA sequence of *Prorocentrum micans*, a dinoflagellate (7).

DISCUSSION

Picoplankton are ubiquitous in oligotrophic ocean waters at concentrations of 10⁵ to 10⁷ cells per ml (2). As is the case in many environments, only a small fraction (<1%) of planktonic cells observed by direct microscopy can be recovered as CFU on laboratory media (8, 10, 23). Study of the phylogenetic diversity of such populations therefore requires techniques that do not rely on cultivation of the resident microorganisms. The approach used here, analysis of rRNA gene sequences retrieved from a shotgun library of naturally occurring picoplankton DNA, seems the most rigorous way to obtain all representatives of the population. The shotgun library, with 10- to 20-kb inserts, is also a source of other sequences associated with the rRNA genes or other genes of interest.

Plankton were harvested by tangential flow filtration using

a 10- μ m-pore-size intake filter and a 0.1- μ m-pore-size tangential flow filter. These size restrictions encompass the picoplankton, organisms between 0.2 and 2 μ m in diameter, the dominant biomass in the oligotrophic ocean (18). The seawater intake for the tangential flow filter was approximately 3 m under the water's surface. Seawater collected at this depth is expected to be representative of the top 90 m of the water column, since depth profiles of dissolved oxygen, temperature, and salinity taken during the collection were uniform to a depth of ca. 90 m, indicative of mixing. Also, since the collection lasted for approximately 3 days and within a 6-mile radius of the ALOHA station, variation from small scale vertical or horizontal patchiness should be minimal.

A rigorous protocol was used to lyse and purify DNA from a potentially diverse collection of organisms. Since there were few intact cells following lysis and no evidence of DNA resistant to restriction endonuclease digestion, we expect minimal selective loss of particular rRNA genes during lysis and the preparation of the size-fractionated DNA. Screening of 3.2×10^4 clones with the mixed-kingdom probe indicated that 0.12% of the clones contain an rDNA insert. This recovery approximates the amount of organismal DNA typically allotted to 16S rRNA genes (0.1 to 0.4%) (13), suggesting that rDNA clones were neither selected for nor excluded during cloning.

The current stage of screening of this picoplankton recombinant library is far from exhaustive. The analysis of 38 rDNA-containing clones was arbitrary; the goal was to obtain some idea of the prominent types of organisms among picoplankton. The fact that many of the sequences were obtained as single instances suggests that the diversity of rRNA types in this recombinant library has not been fully sampled. On the other hand, the phylogenetic distribution of sequences recovered indicates that the picoplankton are dominated by eubacteria: cyanobacteria and α and γ proteobacteria. The relative recoveries of different types of rRNA-containing recombinants does not necessarily reflect in detail the relative abundances of cell types, however, because organisms differ in the number of rRNA genes per cell.

The finding of cyanobacterial rRNA genes among picoplankton was expected, considering the abundance and global distribution of cyanobacteria in ocean waters. All of the cyanobacterial sequences identified are very close relatives of two *Synechococcus* strains in the Woods Hole Culture collection, WH 8103 and WH 7805 (22). Cyanobacterial sequences obtained from picoplankton in the Sargasso Sea, designated SAR 6 and SAR 7 (3), also fall within this closely related group of marine *Synechococcus* strains, consistent with many earlier studies identifying marine *Synechococcus* strains as the most abundant cyanobacteria in the oceans (2, 22). Because of their close evolutionary relationships, the biochemical properties of the cyanobacteria seen here only as rRNA genes are expected to be highly similar to those of the related strains in culture. Certainly their mode of nutrition is expected to be photosynthetic, as is that of all known cyanobacteria.

The other eubacterial rDNA sequences were affiliated with or closely related to two relatedness groups of proteobacteria, the α and γ subdivisions. It is likely that all of these proteobacterial picoplankton are oligotrophic heterotrophs, since proteobacterial photosynthesis is anaerobic. Members of the γ proteobacteria such as *Alteromonas*, *Vibrio*, and *Pseudomonas* spp. are common marine pelagic isolates (2). Some of the isolated rDNA clones are closely related to these genera.

The α subdivision of proteobacteria contains many genera of oligotrophic bacteria that have been isolated from marine waters, including *Caulobacter*, *Hyphomonas*, *Hyphomicrobium*, and *Seliberia* spp. (2). The rDNA sequences from the marine picoplankton that are related to the α proteobacteria are not closely related to these or any other identified α proteobacteria. Some of the α -related sequences are, however, nearly identical to rDNA clones obtained from Sargasso Sea DNA by PCR amplification (3). Their occurrence in both the Pacific (this study) and Atlantic (Sargasso Sea) (3) oceans suggests a global distribution of these currently uncharacterized bacteria.

Studies are under way to assess in more detail the relative distributions of the rDNA sequences in bulk DNA collected from the ALOHA site. Once these methods are optimized, studies of spatial and temporal changes in the picoplankton community structure can be conducted. The use of these approaches will complement existing strategies for studying marine picoplankton and should be generally applicable in microbial ecology.

ACKNOWLEDGMENTS

We thank Dan Distel and John Waterbury for making cyanobacterial 16S rRNA sequences available, and we thank David Karl and colleagues (University of Hawaii) and the crew of the R/V *Moana Wave* for assistance with the collection.

This research was supported by a contract from the Office of Naval Research (N14-87-K-0813) and a grant from the National Institutes of Health (GM34527) to N.R.P.

REFERENCES

1. Atlas, R. M., and R. Bartha. 1981. Microbial ecology. Addison-Wesley Publishing Co., Inc., Reading, Mass.
2. Fogg, G. E. 1986. Picoplankton. Proc. R. Soc. London Ser. B 228:1-30.
3. Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field. 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature (London) 345:60-63.
4. Giovannoni, S. J., E. F. DeLong, G. J. Olsen, and N. R. Pace. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. J. Bacteriol. 170:720-726.
5. Giovannoni, S. J., E. F. DeLong, T. M. Schmidt, and N. R. Pace. 1990. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. 56:2572-2575.
6. Helms, C., M. Y. Graham, J. E. Dutchik, and M. V. Olson. 1985. A new method for purifying lambda DNA from phage lysates. DNA 4:39-49.
7. Herzog, M., and L. Maroteaux. 1986. Dinoflagellate 17S rRNA sequence inferred from the gene sequence: evolutionary implications. Proc. Natl. Acad. Sci. USA 83:8644-8648.
8. Jannasch, H. W., and G. E. Jones. 1959. Bacterial populations in seawater as determined by different methods of enumeration. Limnol. Oceanogr. 4:128-139.
9. Kaiser, K., and N. E. Murray. 1985. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries, p. 1-47. In D. M. Glover (ed.), DNA cloning: a practical approach, vol. 1. IRL Press, Ltd., Oxford.
10. Kogure, K., U. Simidu, and N. Taga. 1979. A tentative direct microscopic method for counting living marine bacteria. Can. J. Microbiol. 25:415-420.
11. Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl. Acad. Sci. USA 82:6955-6959.
12. Olsen, G. J. 1988. Phylogenetic analysis using ribosomal RNA. Methods Enzymol. 164:793-812.
13. Pace, N. R. 1973. Structure and synthesis of the ribosomal ribonucleic acid of prokaryotes. Bacteriol. Rev. 37:562-603.

14. Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1986. The analysis of natural microbial populations by ribosomal RNA sequences. *Adv. Microb. Ecol.* **9**:1–55.
15. Powell, R., J. Neilan, and F. Gannon. 1986. Plaque dot assay. *Nucleic Acids Res.* **14**:1541.
16. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487–491.
17. Sgaramella, V., and H. G. Khorana. 1972. Total synthesis of the structural gene for an alanine transfer RNA from yeast. Enzymatic joining of the chemically synthesized polydeoxynucleotide to form the DNA duplex representing nucleotide sequence 1 to 20. *J. Mol. Biol.* **72**:427–444.
18. Sieburth, J. M., V. Smetacek, and J. Lenz. 1978. Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnol. Oceanogr.* **23**:1256–1263.
19. Stahl, D. A., D. J. Lane, G. J. Olsen, and N. R. Pace. 1984. Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. *Science* **224**:409–411.
20. Stahl, D. A., D. J. Lane, G. J. Olsen, and N. R. Pace. 1985. Characterization of a Yellowstone hot spring microbial community by 5S rRNA sequences. *Appl. Environ. Microbiol.* **45**:1379–1384.
21. Ward, D. M., R. Weller, and M. M. Bateson. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature (London)* **345**:63–65.
22. Waterbury, J. B., S. W. Watson, F. W. Valois, and D. G. Franks. 1986. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. *Can. Bull. Fish. Aquat. Sci.* **214**:71–120.
23. Watson, S. W., T. J. Novitsky, H. L. Quinby, and F. W. Valois. 1977. Determination of bacterial number and biomass in the marine environment. *Appl. Environ. Microbiol.* **33**:940–946.
24. Weller, R., and D. M. Ward. 1989. Selective recovery of 16S rRNA sequences from natural microbial communities in the form of cDNA. *Appl. Environ. Microbiol.* **55**:1818–1822.
25. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
26. Woese, C. R., E. Stackebrandt, T. J. Macke, and G. E. Fox. 1965. A phylogenetic definition of the major eubacterial taxa. *Syst. Appl. Microbiol.* **6**:143–151.